The Annals of Probability
1996, Vol. 24, No. 4, 1926–1952

CLUSTER FORMATION IN A STEPPING-STONE MODEL WITH CONTINUOUS, HIERARCHICALLY STRUCTURED SITES

BY STEVEN N. EVANS[1] AND KLAUS FLEISCHMANN

University of California, Berkeley and Weierstrass Institute for Applied Analysis and Stochastics

A stepping-stone model with site space a continuous, hierarchical group is constructed via duality with a system of (delayed) coalescing "stable" Lévy processes. This model can be understood as a continuum limit of discrete state-space, two-allele, genetics models with hierarchically structured resampling and migration. The existence of a process rescaling limit on suitably related large space and time scales is established and interpreted in terms of the dynamics of cluster formation. This paper was inspired by recent work of Klenke.

1. Introduction and results.

1.1. *Background.* In several physical and biological systems, the phenomenon of *cluster formation* can be observed. One has systems in which spatially dispersed units can be one of two or more possible types. There is a mechanism that attempts to impose local agreement among units, possibly in the face of "noise" that can destroy the agreement but may also spread it. One of the fundamental questions about such systems is the manner in which clusters (i.e., large regions of agreement) grow and interact with each other.

A rather detailed picture on the growth of clusters in the simple *voter model* on the one-dimensional lattice $\mathbb{Z}$ was developed by Arratia [1].

An analogous picture emerged for a certain class of *stepping-stone models* in the work of Klenke [12], Theorem 2. He considered a system of *interacting diffusions* of the Fisher–Wright type with state space $[0, 1]$ indexed by the *countable hierarchical group*

$$(1)\quad \Xi := \{\xi = (\xi^i)_{i\in\mathbb{Z}_-} \in (\mathbb{Z}_N)^{\{\dots,-2,-1\}}\colon\ \xi^i = 0 \text{ for all } i \text{ sufficiently small}\},$$

where $\mathbb{Z}_-$ denotes the negative integers, $\mathbb{Z}_N$ is the *cyclic Abelian group* $\{0, \dots, N-1\}$ of order $N \geq 2$ with the operation of addition modulo N and addition in Ξ is performed coordinatewise. The reason for the nomenclature is that the sets

$$\Xi_k := \{(\xi^i)_{i\in\mathbb{Z}_-}\colon\ 0 = \xi^{-k-1} = \xi^{-k-2} = \cdots\}, \qquad k \in \mathbb{Z}_+$$

Received August 1995; revised January 1996.

[1]Research supported in part by a Presidential Young Investigator Award and an Alfred P. Sloan Foundation Fellowship.

AMS 1991 *subject classifications*. Primary 60K35; secondary 60J60, 60B15, 60J30.

Key words and phrases. Interacting diffusion, stochastic partial differential equation, measure-valued process, stepping-stone model, Fisher–Wright diffusion, cluster formation, clustering, coalescing Lévy process, hierarchical structure, resampling, migration.

($\mathbb{Z}_+$ denotes the nonnegative integers) are finite subgroups of Ξ with $\{0\} = \Xi_0 \subset \Xi_1 \subset \cdots$. Thus, each point of Ξ belongs to a unique coset of Ξ_1 along with $N-1$ other points, each coset of Ξ_1 is contained in a unique coset of Ξ_2 along with $N-1$ other cosets of Ξ_1 and so on.

These models arise as the $M \to \infty$ *diffusion limits* of a class of discrete state-space models in *population genetics* in which the sites represent demes or colonies of M individuals, each possessing one of two possible genotypes. Here the value of the process at a site is the proportion of the colony with a given genotype. These proportions evolve by independent resampling within colonies and migration of individuals between colonies. In this interpretation we can think of the hierarchical structure of Ξ as capturing the idea that colonies are grouped into clans, clans are grouped into villages, villages are grouped into counties and so forth. Consonant with this interpretation, the strength of the migratory flux between two sites is taken to be a function of how far apart the sites are in this hierarchy. We refer the reader to Sawyer and Felsenstein [17] for more discussion of the biology behind the original discrete models (see also Sawyer [16]). We will give a more precise description of the diffusion limits in Section 4.

Klenke [12] showed that if the migration rates coincide with the jump rates of a *"strongly recurrent"* random walk on Ξ, then as time evolves the sites will tend to segregate into increasingly large clusters where the value of the diffusion at the sites in the cluster is close to either 0 or 1; moreover, there is a characteristic rate at which such clusters grow. Although we will not give the precise definition of "strong recurrence" here, it might help the reader's intuition if we remark that the simple random walk on $\mathbb{Z}$ is strongly recurrent, whereas on $\mathbb{Z}^2$ it is not.

Regimes in which the migration rates in interacting diffusions are the jump rates of a recurrent, but not strongly recurrent, random walk were studied by Fleischmann and Greven [10, 11] and Cox, Fleischmann and Greven [4]. The clustering behavior for these latter models is different and rather more subtle. (See [5] and [3] for similar results concerning the related voter model.)

In [12] and [10] two quantitative phenomena are considered as proxies for the somewhat imprecise notion of *cluster formation*. The first is the presence of blocks of sites in which the average value over the block is close to 0 or 1, and the second is the presence of significant "correlations" between sites that are far apart. The latter phenomenon is expressed in terms of the behavior of a sequence of models that is obtained by "thinning out" sites, so that a large number of neighboring sites is replaced by a single representative.

1.2. *Purpose of the paper.* In this paper we consider a class of processes X that also arise as limits of the kind of simple discrete models described above. The difference here is that, loosely put, we pass to a *continuum limit* with the space of sites, so that the smallest geographic units become microscopic entities, rather than remaining as mesoscopic entities as they do in [12] and [10]. Our processes X can be thought of as *infinitesimal cousins* of those in [12]. Instead of $[0, 1]^{\Xi}$, the state space of our processes is the set of Borel functions in $[0,1]^G$, where G is the *hierarchical group* of all semiinfinite sequences,

$$G := \{ g = (g^i)_{i \in \mathbb{Z}} \in (\mathbb{Z}_N)^{\mathbb{Z}}:\ g^i = 0 \text{ for all } i \text{ sufficiently small} \} \tag{2}$$

(again with coordinatewise addition), a group that can be topologized as a *nondiscrete*, locally compact, totally disconnected group. Our processes are natural stochastic partial differential equation (SPDE) analogs [see (8) below] of the infinite system of stochastic differential equations considered in [12] and arise as limits of the latter processes (cf. the proof of Theorem 3 in Section 4 below). In particular, the "drift part" of the SPDE is determined by the jump rates of a "stable" Lévy process on G.

We also remark that our processes are essentially particular examples of the continuum stepping-stone models considered in [19].

As well as being of interest in their own right, a significant advantage of our models is that they exhibit the same sort of cluster formation dynamics as the models in [12], but these phenomena can be more easily described and understood in our setting. More precisely, our models can be rescaled at suitably related large time and space scales to obtain limiting processes that also have the Borel functions from G to $[0,1]$ as their state space. Results about the formation of clusters in our original models can then be rephrased as easily proven facts about the microscopic and macroscopic spatial structure at fixed times of these scaling limits. In particular, there is no need to resort to "artifices" such as thinning or block averaging. These latter transformations can be seen as partial substitutes for spatial rescalings that are unavailable in models with a discrete collection of sites. Moreover, our point of view enables us to study the evolution of all the clusters and not just the cluster containing the origin.

A model analogous to ours was considered in Mueller and Tribe [15] with G replaced by $\mathbb{R}$ and the Lévy process that describes the migration replaced by Brownian motion. This analog arises as a suitable scaling limit of a long-range voter process on $\mathbb{Z}$. It appears that it is possible to construct a sequence of long-range voter process-like particle systems on Ξ that can be rescaled in the manner of [15] to converge to our process, but we do not pursue this matter in the present paper.

1.3. *The site set G.* Before we can describe more precisely the process we wish to consider, we need to make a few simple remarks about the structure of the group G of (2). General discussions of the structure of totally disconnected, locally compact, Abelian groups may be found in [21] or [9]. Via

$$|g| := N^{-k}, \quad \text{where } g \in G \text{ and } k := \inf \{ i \in \mathbb{Z}:\ g^i \neq 0 \}, \tag{3}$$

we introduce a translation-invariant *ultrametric* on G, that is, a translation-invariant metric satisfying

$$|g - g'| \le |g| \vee |g'|, \qquad g, g' \in G.$$

With this metric, G is a *nondiscrete, locally compact, totally disconnected Abelian group with countable base*. Note that the balls

$$G_k := \{g \in G\colon\ |g| \le N^{-k}\}, \qquad k \in \mathbb{Z}, \tag{4}$$

are compact-open subgroups of G satisfying $G_k \supset G_{k+1}$, and that

$$|g| = N^{-k} \quad \text{if and only if } g \in G_k \backslash G_{k+1}. \tag{5}$$

Denote by $l(dg) = dg$ the *Haar measure* on G, normalized so that $l(G_0) = 1$. That is,

$$l \text{ assigns the mass } N^{-k} \text{ to } G_k, \qquad k \in \mathbb{Z}, \tag{6}$$

and, conditioned on G_k, it has i.i.d. coordinates g^i for $i \ge k$, "uniformly" distributed on $\mathbb{Z}_N$. In particular,

$$l(G_k \backslash G_{k+1}) = N^{-k}(1 - N^{-1}), \qquad k \in \mathbb{Z}. \tag{7}$$

1.4. *Description of the model: existence of X.* *Formally*, the process we wish to consider is the process X which has as its state space the space of Borel maps from G into $[0, 1]$ and solves the stochastic partial differential equation

$$\begin{aligned} dX_t(g) \otimes dg &= \sqrt{af(X_t(g))}\, w(dt \otimes dg) \\ &\quad + \left\{ \int_G \nu(dg')[X_t(g + g') - X_t(g)] \right\} dt \otimes dg, \end{aligned} \tag{8}$$

$t > 0$, $g \in G$. Here $w(dt \otimes dg)$ is *time–space white noise* with $\mathbb{P}[\{w(A \times B)\}^2] = m(A)l(B)$ for Borel sets $A \subset \mathbb{R}_+$ and $B \subset G$, with m being the Lebesgue measure on $\mathbb{R}_+$ and l being the Haar measure on G defined in Section 1.3. Further, f is the standard *Fisher–Wright diffusion coefficient*

$$f(r) := r(1 - r), \qquad 0 \le r \le 1, \tag{9}$$

and ν is the *Lévy measure*

$$\nu(dg) := b\,|g|^{-\alpha-1}\, dg, \tag{10}$$

where $1 < \alpha < \infty$ and a, $b > 0$ are fixed constants, called the *Lévy index, diffusion constant* and *Lévy constant*, respectively.

The reader familiar with the Fleming–Viot process may notice some similarity between that process and ours. The difference is that in our process resampling only occurs within the individuals at each site, rather than across the whole population.

An existence and uniqueness theorem for this type of SPDE is stated with a briefly sketched proof as Theorem 5.1(ii) in [19]. As we wish to consider rescaling limits of X that do not appear to be solutions to SDPE's, it will be more convenient for us to define the process X by describing it as a Feller process with an explicitly given semigroup.

The key to such a description is the observation in [19] that a solution to (8) is *dual*, via moment functions, to a *(delayed) coalescing Lévy process*. That is, the dual can be thought of as a finite system of unlabeled particles that move independently in G as "stable" Lévy processes with Lévy measure ν of (10), but, additionally, each pair of colliding particles coalesces to a single particle at rate a times their collision local time (i.e., the local time at 0 of the difference of their positions.)

This description of the dual is not quite what we will use. Instead, we will consider a slightly enhanced model in which we have a finite system of particles *labeled* by $\{1, \ldots, n\}$, $n \in \mathbb{N}$, that move independently in G as Lévy processes with Lévy measure ν, but, additionally, each particle can be killed and sent to an adjoined *cemetery* state $\dagger$ at rate a times the total of the collision local times between the particle and the other living particles with smaller labels. We will denote this latter process by $(\vartheta,\ \mathbf{P}^{\mathbf{g}}) = (\vartheta,\ \mathbf{P}^{\mathbf{g}}_{a,b})$ when the initial state is $\mathbf{g} = (g_1, \ldots, g_n) \in G^n_\dagger := (G \cup \{\dagger\})^n$. We call it the *(delayed) coalescing Lévy process*. A fuller description is given in Section 3.1.

As a final preliminary, we need to say something about the *state space* $\mathbf{B}$ that we will use for our process X. Let $\mathbf{B}$ denote the set of equivalence classes of Borel functions from G into $[0, 1]$, where we declare that two functions are equivalent if they are equal l-a.e. (recall that l is the Haar measure on G). We can associate $x \in \mathbf{B}$ with the Radon measure $x(g)\,dg$ on G. Via this identification, we can think of $\mathbf{B}$ as a closed subset of the space of all Radon measures on G endowed with the vague topology. (In this sense, the process X to be constructed can be understood as a *measure-valued diffusion*.)

Alternatively, we can regard $\mathbf{B}$ as a closed subset of $L^\infty = L^\infty(G, l)$, furnished with its weak* topology as the dual of $L^1 = L^1(G, l)$.

These two relative topologies on $\mathbf{B}$ coincide. As both are metrizable, to see this it suffices to show that, for $x_0,\ x_1, \ldots \in \mathbf{B}$,

$$\int_G dg\, x_j(g)\, \varphi(g) \to \int_G dg\, x_0(g)\, \varphi(g) \quad \text{as } j \to \infty$$

holds for all φ in the set $C_c(G)$ of all continuous functions φ on G with compact support, if and only if it holds for all $\varphi \in L^1$. But this is immediately clear since $C_c(G)$ is dense in L^1 and the x_j are uniformly bounded.

By Corollary 5.4.3 of Dunford and Schwartz [6], this $\mathbf{B}$ is a *compact metrizable space*.

DEFINITION 1 (Product brackets). If x is a function defined on G, and $n \in \mathbb{N}$, we set

$$[x,\ \mathbf{g}] := \prod_i \mathbf{1}\{g_i \neq \dagger\}\, x(g_i), \qquad \mathbf{g} = (g_1, \ldots, g_n) \in G^n_\dagger\,.$$

LEMMA 2 (Weight functions). *For $n \in \mathbb{N}$ and $\varphi \in L^1(G^n, l^n)$, the function I_n^φ: $\mathbf{B} \to \mathbb{R}$ defined by*

$$I_n^\varphi(x) := \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, [x,\, \mathbf{g}], \qquad x \in \mathbf{B}, \tag{11}$$

is continuous.

PROOF. For φ of the form $\varphi(\mathbf{g}) = \varphi_1(g_1), \dots, \varphi_n(g_n)$ with $\varphi_i \in C_c(G)$, the claim is immediate. The general statement follows once we note that linear combinations of such functions are dense in $L^1(G^n,\, l^n)$, and if $\{\varphi_j\}_{j=1}^\infty$ is a sequence in $L^1(G^n,\, l^n)$ that converges to φ, then $\{I_n^{\varphi_j}\}_{j=1}^\infty$ converges uniformly to I_n^φ. $\square$

Our first result is the following existence theorem, the proof of which is postponed to Section 4.

THEOREM 3 (Existence of the stepping-stone process X). *For $a, b > 0$ there exists a unique strongly continuous semigroup $\mathsf{S} = \mathsf{S}_{a,b} = \{\mathsf{S}_{a,b}(t)\colon\ t \geq 0\}$ of Markov linear operators $\mathsf{S}_{a,b}(t)\colon\ C(\mathbf{B}) \to C(\mathbf{B})$ (i.e., a Feller semigroup) such that*

$$\mathsf{S}_{a,b}(t) I_n^\varphi(x) = \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, \mathbf{P}_{a,b}^{\mathbf{g}}[x,\, \vartheta_t], \tag{12}$$

$t \geq 0$, $n \geq 1$, $\varphi \in L^1(G^n,\, l^n)$, $x \in \mathbf{B}$. *Moreover, there is a Hunt process $(X,\, \mathbb{P}_{a,b}^x)$ on $\mathbf{B}$ with continuous sample paths and semigroup $\mathsf{S}_{a,b}$.*

This $(X,\, \mathbb{P}_{a,b}^x)$ is our *stepping-stone process* with diffusion constant a and Lévy constant b.

1.5. *The limiting cluster process Y.* In order to describe the large-scale space–time properties of X, we need to introduce another $\mathbf{B}$-valued process. By analogy with the definition of the coalescing Lévy process ϑ, we can consider an *instantaneously* coalescing Lévy process. This is a finite system of labeled particles that move independently in G as Lévy processes with Lévy measure ν, but, additionally, when two particles collide the one with the higher label is sent to the cemetery † instantaneously. The state space of this process is the set $\check{G}_\dagger^n$ consisting of n-tuples $(g_1, \dots, g_n) \in G_\dagger^n$ for which $g_i = g_j \neq \dagger$ does not hold for $1 \leq i \neq j \leq n$. We will denote this *instantaneously coalescing Lévy process* by $(\eta,\, \mathbf{Q}_b^{\mathbf{g}})$ when the initial state is $\mathbf{g} = (g_1, \dots, g_n) \in \check{G}_\dagger^n$. A fuller description is given in Section 3.2.

Next we state the existence of the limiting cluster process Y which is proved in Section 4.

THEOREM 4 (Existence of the cluster process Y). *For $b > 0$ there exists a unique strongly continuous semigroup $\mathsf{T} = \mathsf{T}_b = \{\mathsf{T}_b(t)\colon\ t \geq 0\}$ of Markov* *linear operators* $\mathsf{T}_b(t)$: $C(\mathbf{B}) \to C(\mathbf{B})$ (*i.e., a Feller semigroup*) *such that*

$$\mathsf{T}_b(t)I_n^{\varphi}(x) = \int_{\check{G}^n} d\mathbf{g}\ \varphi(\mathbf{g})\mathbf{Q}_b^{\mathbf{g}}[x,\ \eta_t], \tag{13}$$

$t \geq 0$, $n \geq 1$, $\varphi \in L^1(G^n,\ l^n)$, $x \in \mathbf{B}$. *For each* $F \in C(\mathbf{B})$,

$$\lim_{a\to\infty} \mathsf{S}_{a,\,b}(t)F = \mathsf{T}_b(t)F.$$

Moreover, there is a Hunt process $(Y,\ \mathbb{Q}_b^y)$ *on* $\mathbf{B}$ *with continuous sample paths and semigroup* T_b.

We call $(Y,\ \mathbb{Q}_b^y)$ the *cluster process* of X with Lévy constant b. Intuitively, as $a \to \infty$ the diffusion part in X speeds up without limit which should imply that each component $X(g)$ of X will be *trapped* at the boundary $\{0, 1\}$ of the interval $[0, 1]$. [See Theorem 6(iv) below.]

1.6. *Scaling properties of X and Y.* Let σ: $G \to G$ denote the *shrinking automorphism* which moves all the coordinates of a point $g \in G$ by one step to the right, so that $|\sigma g| = N^{-1}|g|$. Using this with a slight abuse of notation, define σ: $\mathbf{B} \to \mathbf{B}$ by $\sigma x = x \circ \sigma$, $x \in \mathbf{B}$, to get an associated bijection on $\mathbf{B}$. With another slight abuse of notation, we will also let σ denote the map from the space of probability measures μ on $\mathbf{B}$ into itself that is given by $\int(\sigma\mu)(dx)\, F(x) = \int \mu(dx)\, F(\sigma x)$ for F a bounded Borel function on $\mathbf{B}$.

Finally, we introduce a group of *space–time scaling transformations* $\Theta = \{\Theta_{m,\,s}\colon m, s \in \mathbb{Z}\}$ on $D(\mathbb{R}_+,\ \mathbf{B})$ (the Skorohod space of cadlag paths from $\mathbb{R}_+$ to $\mathbf{B}$) by

$$(\Theta_{m,\,s}z)_t := \sigma^{-m} z_{N^{\alpha s}t}, \qquad z \in D(\mathbb{R}_+,\ \mathbf{B}),\ t \geq 0. \tag{14}$$

(Recall that α is the index of our "stable" Lévy process.)

It turns out that $\Theta_{m,\,s}X$ (resp. $\Theta_{m,\,s}Y$) is the same sort of process as X (resp. Y).

PROPOSITION 5 (Scaling properties). *Consider m, $s \in \mathbb{Z}$ and a law μ on $\mathbf{B}$. The distribution of $\Theta_{m,\,s}X$ under $\mathbb{P}_{a,\,b}^{\mu}$ (resp. $\Theta_{m,\,s}Y$ under $\mathbb{Q}_b^{\mu}$) is that of X under $\mathbb{P}^{\sigma^{-m}\mu}_{N^{\alpha s - m}a,\ N^{\alpha(s-m)}b}$ (resp. Y under $\mathbb{Q}^{\sigma^{-m}\mu}_{N^{\alpha(s-m)}b}$).*

1.7. *Main result*: *cluster formation of X.* Let $\mathbf{B}_{\{0,\,1\}}$ denote the Borel subset of $\mathbf{B}$ consisting of equivalence classes with a representative that takes values in the set $\{0, 1\}$. Now we have gathered together all the ingredients to formulate our *main result*.

THEOREM 6 (Cluster formation). *Suppose that μ is a shift-invariant and ergodic probability measure on $\mathbf{B}$ with intensity $\theta \in (0, 1)$:*

$$\int \mu(dx) \int dg\, f(g)\, x(g) = \theta \int dg\, f(g), \qquad f \in \mathbf{B}. \tag{15}$$

Then the following statements hold:

(i) *The law of* $\Theta_{m,m}X$ *under* $\mathbb{P}^{\mu}_{a,b}$ *converges to the law of* Y *under* $\mathbb{Q}^{\theta\mathbf{1}}_{b}$ *as* $m \to \infty$.

(ii) *The law of the* $D(\mathbb{R}_+, \mathbf{B})^{\mathbb{Z}}$-*valued random variable* $(\Theta_{m-j,m}X)_{j\in\mathbb{Z}}$ *under* $\mathbb{P}^{\mu}_{a,b}$ *converges to the law of* $\big(\Theta_{-j,0}Y\big)_{j\in\mathbb{Z}}$ *under* $\mathbb{Q}^{\theta\mathbf{1}}_{b}$ *as* $m \to \infty$.

(iii) *For* $t > 0$, *the law of* $(\Theta_{-j,0}Y)_t$ *under* $\mathbb{Q}^{\theta\mathbf{1}}_{b}$ *converges to the two-point mixture*

$$\theta\delta_{\mathbf{1}} + (1-\theta)\delta_{\mathbf{0}} \quad \textit{as } j \to \infty,$$

and to the point mass

$$\delta_{\theta\mathbf{1}} \quad \textit{as } j \to -\infty.$$

(iv) *For* $t > 0$ *fixed*, Y_t *belongs to* $\mathbf{B}_{\{0,1\}}$, $\mathbb{Q}^{\theta\mathbf{1}}_{b}$-*a.s.*

Thus, if we observe X on a suitable collection of large space–time scales, then we see the cluster process Y in the limit. Varying the relationship between the growth of the time and space scales when taking the limit is equivalent to observing Y on different space scales. If we observe Y_t on a microscopic scale, then we find ourselves in the middle of a cluster of 0's or 1's. On the other hand, if we observe Y_t macroscopically, the clusters of 0's and 1's will be averaged, leading to a constant density θ.

The sequence of block-averaging limits studied in [12] corresponds in our setting to the sequence of random variables

$$\left(\int_{G_0} dg\, \Theta_{-j,0}Y_t(g)\right)_{j\in\mathbb{Z}} = \left(N^j \int_{G_j} dg\, Y_t(g)\right)_{j\in\mathbb{Z}}.$$

It is immediate from the spatial stationarity of Y_t that this sequence is a *martingale*, a phenomenon noted in [12].

As an aside, we note that the cluster state Y_t is certainly *random* because of the randomness of the $j \to \infty$ limit. Moreover, the distribution of Y_t cannot be just such a two-point mixture because then the $j \to -\infty$ limit would not hold.

Finally, we remark that a fortiori we have for $t > 0$ and a sequence $(c_j)_{j\in\mathbb{N}}$ of positive integers that as $j \to \infty$ the distribution of $X_{N^{\alpha j}t}(\sigma^{-c_j}\,\cdot)$ converges to the mixture

$$\theta\delta_{\mathbf{1}} + (1-\theta)\delta_{\mathbf{0}} \quad \text{if } c_j/j \to 0,$$

and to the point mass

$$\delta_{\theta\mathbf{1}} \quad \text{if } c_j/j \to +\infty.$$

2. Stable Lévy process of index α. The purpose of this section is to introduce the underlying migration process, a particular Lévy process Z on G, a little more formally and collect some of its properties.

2.1. *More about G.* The basic facts about totally disconnected, locally compact, Abelian groups needed here may be found in [21].

For $k \in \mathbb{Z}$, consider the *quotient group* G/G_k and the related *quotient map* $\pi_k\colon G \to G/G_k$. Since $|\cdot|$ defined in (3) is constant on the cosets of G_k other than G_k itself, in G/G_k we get a translation-invariant *ultrametric* via

$$|\bar{g}| := \begin{cases} |g|, \text{ if } \pi_k g = \bar{g} \neq 0, \\ 0, \quad \text{if } \bar{g} = 0, \end{cases} \qquad \bar{g} \in G/G_k. \tag{16}$$

The balls

$$(G/G_k)_j := \{\bar{g} \in G/G_k \colon |\bar{g}| \leq N^{-k+j}\}, \qquad j \geq 0, \tag{17}$$

are finite subgroups of G/G_k. In particular, $(G/G_k)_0 = \{0\}$, and $(G/G_k)_1$ is isomorphic to the cyclic group $\mathbb{Z}_N$. Note also that if Ξ is the countable hierarchical group defined in (1), then

$$\text{for all } k \in \mathbb{Z}, \text{ the quotient group } G/G_k \text{ is isomorphic to } \Xi. \tag{18}$$

Recall that $G_\dagger = G \cup \{\dagger\}$, where $\dagger$ is adjoined as an isolated cemetery point. Adjoin to G/G_k an isolated cemetery point that we will also denote by the symbol $\dagger$. Extend the quotient maps π_k to $G_\dagger$ by setting $\pi_k(\dagger) := \dagger$.

We also need the *dual group* G^* of G. It can be defined as G in (2) except we reflect the index $j \in \mathbb{Z}$ to $-j$. That is, the elements h of G^* have the 0's at the right end. Set

$$|h| := N^k \text{ where } h \in G^* \text{ and } k := \sup\{j \in \mathbb{Z} \colon h^j \neq 0\} + 1,$$

as well as

$$G_k^* := \{h \in G^* \colon |h| \leq N^k\}, \qquad k \in \mathbb{Z}. \tag{19}$$

Then

$$|h| = N^k \quad \text{if and only if } h \in G_k^* \backslash G_{k-1}^*. \tag{20}$$

The *pairing* $\langle g, h\rangle$ between G and G^* is just given by

$$\langle g, h\rangle := \exp\left[\frac{2\pi i}{N} \sum_{j \in \mathbb{Z}} g^j h^j\right], \qquad g \in G,\ h \in G^*, \tag{21}$$

where for the $g^j, h^j \in \mathbb{Z}_N = \{0, \ldots, N-1\}$ the product $g^j h^j$ is defined by the usual multiplication in $\mathbb{Z}$. Note that

$$G_k^* = \{h \in G^* \colon \langle g, h\rangle = 1,\ \forall g \in G_k\}, \qquad k \in \mathbb{Z}; \tag{22}$$

that is, G_k^* is the *annihilator* of G_k. If μ is a finite measure on G, we define the *Fourier transform* $\widehat{\mu}$ of μ by

$$\widehat{\mu}(h) := \int_G \mu(dg)\,\langle g, h\rangle, \qquad h \in G^*. \tag{23}$$

Write also $\widehat{\varphi}$ instead of $\widehat{\mu}$ if $\mu(dg) = \varphi(g)\,dg$, that is, if φ is the density function of μ.

EXAMPLE 7 (Fourier transforms of some indicators). The Fourier transform of the indicator function $\mathbf{1}_{G_k}$ of the compact-open subgroup G_k is given by

$$\widehat{\mathbf{1}_{G_k}} = N^{-k}\mathbf{1}_{G_k^*}, \qquad k \in \mathbb{Z}. \tag{24}$$

In fact, if $h \in G_k^*$, then $\langle g, h\rangle = 1$ [recall (22)], and (24) follows for such h from $l(G_k) = N^{-k}$. On the other hand, if $h \notin G_k^*$, then there is a $j \geq k$ such that the jth coordinate h^j of h is different from 0. However, g^j is "uniform" on $\mathbb{Z}_N$, and

$$\sum_{g^j=0}^{N-1} \exp\left[\frac{2\pi i}{N}\, g^j h^j\right] = 0, \qquad h^j = 1, \dots, N-1,$$

which implies that (24) is also true for those h.

LEMMA 8 (Approximate identity). *If* $\varphi \in L^1(G^n, l^n)$, $n \in \mathbb{N}$, *then*

$$\lim_{k\to\infty} N^{nk} \int_{G_k^n} d\mathbf{h}\, \varphi(\mathbf{g}+\mathbf{h}) = \varphi(\mathbf{g}) \qquad l^n\text{-a.e. and in } L^1(G^n, l^n).$$

PROOF. It suffices to consider the case when φ is supported on G_{-r}^n for some $r \in \mathbb{N}$. Then for $k \in \mathbb{N}$ the function $g \mapsto N^{-nr}N^{nk}\int_{G_k^n} d\mathbf{h}\,\varphi(\mathbf{g}+\mathbf{h})$ is just the conditional expectation of φ under the probability measure $N^{-nr}l^n(\cdot \cap G_{-r}^n)$ given the σ-field generated by the cosets of G_k^n, and the result follows from the martingale convergence theorem. □

2.2. *Stable Lévy process Z on G.* Let $Z := \{Z_t : t \geq 0\}$ denote the "*stable*" *Lévy process* on G with Lévy measure ν as defined in (10) with the index $1 < \alpha < \infty$ fixed and $b > 0$. That is, Z is a cadlag jump process with stationary independent increments, where a jump with value g occurs in the interval dt with rate $\nu(dg)\,dt$. Consequently, by (5) and (7),

(25) Z makes a jump of size $|g| = N^{-k}$ at rate $bN^{k\alpha}(1 - N^{-1})$, $\qquad k \in \mathbb{Z}$.

Note that ν is indeed a Lévy measure by the finiteness of

$$\begin{aligned}\nu(G\backslash G_k) &= \sum_{j=0}^{\infty} \nu\big(G_{k-1-j}\backslash G_{k-j}\big)\\ &= bN^{k\alpha}\frac{1-N^{-1}}{N^{\alpha}-1}, \qquad k \in \mathbb{Z}.\end{aligned} \tag{26}$$

For more about processes such as Z, we refer to [9], Section 2. Denote by $P^g = P_b^g$ the law of Z starting at $Z_0 = g$. In what follows, we simply call (Z, P_b^g) the *Lévy process* (with Lévy constant b).

Next we want to calculate the characteristic function of Z_t [recall (21)].

LEMMA 9 (Characteristic function of Z_t). *For each $t > 0$, the characteristic function of Z_t under P_b^0 is given by*

$$P_b^0\langle Z_t, g^*\rangle = \exp\big[-c\,b\,t\,|g^*|^{\alpha}\big], \qquad t \geq 0,\ g^* \in G^*, \tag{27}$$

with the constant

$$c = c_{N,\alpha} := \frac{1 - N^{-\alpha-1}}{N^{\alpha} - 1}.$$

PROOF. Since Z is a Lévy process with Lévy measure ν, for $t > 0$ fixed we have

$$P_b^0\langle Z_t, g^*\rangle = \exp\left[-t\int_G \nu(dg)\big(1 - \langle g, g^*\rangle\big)\right], \qquad g^* \in G^*;$$

see [9], Proposition 1 (send there $N \to \infty$). It remains to show that

$$\int \nu(dg)\big(1 - \langle g, g^*\rangle\big) = c_{N,\alpha}\, b\,|g^*|^{\alpha}, \qquad g^* \neq 0. \tag{28}$$

Decompose the left-hand side into a sum of the contributions from each "annulus" $G_j \backslash G_{j+1}$ and apply (5) to conclude that the left-hand side is

$$b \sum_{j\in\mathbb{Z}} N^{(\alpha+1)j} \int_{G_j\backslash G_{j+1}} dg\big(1 - \langle g, g^*\rangle\big).$$

Using (7) and (24), we see that this is equal to

$$b \sum_{j\in\mathbb{Z}} N^{\alpha j}\big[\mathbf{1}_{G^*\backslash G_j^*} - N^{-1}\mathbf{1}_{G^*\backslash G_{j+1}^*}\big](g^*).$$

Assume now that $|g^*| = N^k$, $k \in \mathbb{Z}$. Then the latter expression coincides with

$$b\sum_{j<k} N^{\alpha j} - bN^{-1}\sum_{j<k-1} N^{\alpha j} = bN^{\alpha(k-1)} + b(1 - N^{-1})\sum_{j<k-1} N^{\alpha j}.$$

But this equals the right-hand side of (28), completing the proof. □

Recall that σ denotes the shrinking automorphism defined in the beginning of Section 1.6. Sometimes we write $Z(t)$ instead of Z_t.

COROLLARY 10 (Scaling for Z). *For $m \in \mathbb{Z}$, $s \in \mathbb{R}$ and $g \in G$, the distribution of the process $\sigma^m Z(N^{\alpha s}\cdot)$ under P_b^g is that of the process Z under $P^{\sigma^m g}_{N^{\alpha(s-m)}b}$.*

PROOF. Using the Lévy property, without loss of generality we may set $g = 0$. It suffices by the Markov property and a simple induction argument to show that for all $t \geq 0$ the distribution of the random variable $\sigma^m Z(N^{\alpha s}t)$ under P_b^0 is that of the random variable $Z(t)$ under $P^0_{N^{\alpha(s-m)}b}$.

Let $\sigma^*\colon G^* \to G^*$ denote the "*adjoint*" *shrinking automorphism* that moves every coordinate of g^* to the left, so that $|\sigma^* g^*| = N^{-1}|g^*|$ and $\langle g, \sigma^* g^*\rangle = \langle \sigma g, g^*\rangle$ for $g^* \in G^*$ and $g \in G$. Then the characteristic function of $\sigma^m Z(N^{\alpha s}t)$ under P_b^0 is given by

$$P_b^0\langle Z(N^{\alpha s}t), (\sigma^*)^m g^*\rangle = \exp\big[-c\, bN^{\alpha(s-m)}t|g^*|^{\alpha}\big],$$

where we used Lemma 9. Applying that lemma again, the claim follows. □

COROLLARY 11 (Transition density of Z). *The Lévy process Z has a jointly continuous transition density $p = \{p_t(g): t > 0, g \in G\}$ with respect to the Haar measure l that is strictly positive and uniformly bounded on each set of the form $[\varepsilon, \infty) \times G$, where $\varepsilon > 0$.*

PROOF. It follows from Lemma 9, the characterization (20) and Example 7 that

$$P_b^0\langle Z_t, g^*\rangle = \sum_{k\in\mathbb{Z}} c_k(t)\mathbf{1}_{G_k^*}(g^*) = \sum_{k\in\mathbb{Z}} c_k(t)\, N^k\widehat{\mathbf{1}_{G_k}}(g^*),$$

$g^* \in G^*$, where we set

$$c_k(t) := \exp\bigl[-c\,b\,t\,N^{k\alpha}\bigr] - \exp\bigl[-c\,b\,t\,N^{(k+1)\alpha}\bigr].$$

Thus, $P_b^0\{Z_t \in dg\} = p_t(g)\,l(dg)$, where $p_t(g) := \sum_{k\in\mathbb{Z}} c_k(t)N^k\mathbf{1}_{G_k}$. It is immediate that the transition density p has the desired properties. □

COROLLARY 12 (Equivalence of restricted laws). *For $\varepsilon > 0$ and $g, h \in G$, the restrictions of P_b^g and P_b^h to the sub-σ-field $\sigma\{Z_t\colon\ t \geq \varepsilon\}$ are equivalent.*

PROOF. This is immediate from Corollary 11 and the Markov property. □

2.3. *Local time Λ for Z.* Later on we will make use of the following fact.

PROPOSITION 13 (Local time of Z). *For each $g \in G$ we have P_b^g-a.s. that there is a jointly continuous local time $(t, h) \mapsto \Lambda(t, h)$, $(t, h) \in \mathbb{R}_+ \times G$, such that*

$$\int_0^t ds\, f(X_s) = \int_G dh\, \Lambda(t, h) f(h)$$

for all bounded Borel functions f and all $t \geq 0$. In particular, P_b^g-a.s.,

$$\Lambda(t, h) = \lim_{k\to\infty} N^k \int_0^t ds\, \mathbf{1}\bigl\{|Z_s - h| \leq N^{-k}\bigr\}, \tag{29}$$

uniformly for (t, h) in compact subsets of $\mathbb{R}_+ \times G$. Moreover, for fixed $h \in G$,

$$\inf\bigl\{t > 0: Z_t = h\bigr\} = \inf\bigl\{t > 0: \Lambda(t, h) > 0\bigr\} < \infty, \qquad P_b^g\text{-a.s.} \tag{30}$$

PROOF.
Step 1. *Existence.* For $\lambda > 0$, write

$$u^\lambda(g) := \int_0^\infty dt\, \mathrm{e}^{-\lambda t}\, p_t(g), \qquad g \in G,$$

for the λ-*potential density* of Z. By (27), its Fourier transform $\widehat{u^\lambda}$ is

$$\widehat{u^\lambda}(g^*) = \int_0^\infty dt \exp\bigl[-\lambda t - c\,b\,t\,|g^*|^\alpha\bigr] = \frac{1}{\lambda + c\,b\,|g^*|^\alpha} = \sum_{k\in\mathbb{Z}} d_k \mathbf{1}_{G_k^*}(g^*),$$

$g^* \in G^*$, where we set

$$d_k := \frac{1}{\lambda + c\,b\,N^{k\alpha}} - \frac{1}{\lambda + c\,b\,N^{(k+1)\alpha}} \geq 0, \qquad k \in \mathbb{Z}.$$

Note that

$$d_k \text{ is of order } N^{-|k|\alpha} \quad \text{as } |k| \to \infty. \tag{31}$$

In particular, d_k is summable in $k \in \mathbb{Z}$. Now, by (24), $\mathbf{1}_{G_k^*}$ is the Fourier transform of $N^k \mathbf{1}_{G_k}$, for each $k \in \mathbb{Z}$. Hence,

$$0 \leq u^\lambda(g) = \sum_{k \in \mathbb{Z}} d_k\, N^k\, \mathbf{1}_{G_k}(g) \leq u^\lambda(0) < \infty, \qquad g \in G. \tag{32}$$

Therefore, u^λ is a bounded continuous function on G, and from

$$u^\lambda(0) - u^\lambda(g) = \sum_k \mathbf{1}\{N^{-k} < |g|\}\, d_k\, N^k$$

and (31) we conclude

$$\lim_{0 \neq g \to 0} \frac{u^\lambda(0) - u^\lambda(g)}{|g|^{\alpha - 1}} = k_{b,\,N,\,\alpha,\,\lambda}$$

for some constant $k_{b,\,N,\,\alpha,\,\lambda} \in (0, \infty)$. As in the proof of Lemma 7.2 of [8], we can check Dudley's metric entropy condition to conclude that there is a version of the centered, stationary *Gaussian process* on G with covariance kernel $(g, g') \mapsto u^1(g' - g)$ that has continuous sample paths. The *existence* of a continuous local time Λ then follows from Theorem 1 of Marcus and Rosen [14]. The limit relation (29) follows from general theory.

Step 2. *Stopping time identity*. Fix $g, h \in G$. Write V_h and W_h, respectively, for the stopping times on the left- and right-hand sides of (30). Observe that the right continuity of Z implies that $Z_{V_h} = h$ on the event $\{V_h < \infty\}$, P_b^g-a.s. Similarly, $Z_{W_h} = h$ on the event $\{W_h < \infty\}$, and $V_h \leq W_h$, P_b^g-a.s.

Let us first show that

$$P_b^{g'}\{W_{h'} < \infty\} > 0 \quad \text{for all } g', h' \in G. \tag{33}$$

By Fubini's theorem, $\int de\, P_b^0 \Lambda(t, e) = P_b^0 \int de\, \Lambda(t, e) = t$ for all $t \geq 0$, and so $P_b^0\{W_e < \infty\} > 0$ for some $e \in G$. By Corollary 12 we get $P_b^f\{W_e < \infty\} > 0$ for all $f \in G$, and combining this with the Lévy property establishes (33).

Let us now show that $V_h = W_h$, P_b^g-a.s. It suffices by applying the strong Markov property at time V_h on the event $\{V_h < \infty\}$ to show that

$$P_b^h\{W_h = 0\} = 1, \tag{34}$$

but this follows by applying the strong Markov property at time W_h on the positive probability event $\{W_h < \infty\}$ [recall (33)].

We are thus left with showing that $P_b^g\{V_h < \infty\} = 1$. By (34) we know that the random set $\{t > 0\colon Z_t = h\}$ is nonempty P_b^h-a.s. We have from Corollary 10 that, under P_b^0, the distribution of Z is the same as that of the process $\sigma^k Z(N^{\alpha k}\cdot)$ for all $k \in \mathbb{Z}$. Consequently, by the Lévy property, under P_b^h, the distribution of the random set $\{t: Z(t) = h\}$ is the same as that of $\{t: Z(N^{\alpha k} t) = h\} = N^{-\alpha k}\{t: Z(t) = h\}$. Sending $k \to -\infty$, we see that the random set $\{t: Z(t) = h\}$ is unbounded P_b^h-a.s. Thus, by Corollary 12, the random set $\{t: Z(t) = h\}$ is unbounded P_b^g-a.s.; hence, $P_b^g\{V_h < \infty\} = 1$ follows. □

COROLLARY 14 (Collision local time). *Let $(\tilde{Z}_t, \tilde{P}_b^{\tilde{g}})$ be a copy of (Z_t, P_b^g). Then for $g, \tilde{g} \in G$ we have $P_b^g \times \tilde{P}_b^{\tilde{g}}$-a.s. that the limit*

$$L(t) := \lim_{k\to\infty} N^k \int_0^t ds\, \mathbf{1}\{|Z_s - \tilde{Z}_s| \le N^{-k}\}$$

exists uniformly on compact subsets of $\mathbb{R}_+$, and this collision local time $L(t)$ of Z and $\tilde{Z}$ is continuous in t. Moreover,

$$\inf\{t > 0: Z_t = \tilde{Z}_t\} = \inf\{t > 0: L(t) > 0\} < \infty, \qquad P_b^g \times \tilde{P}_b^{\tilde{g}}\text{-a.s.} \tag{35}$$

PROOF. This is immediate from Proposition 13 and the observation that the law of $Z - \tilde{Z}$ under $P_b^g \times \tilde{P}_b^{\tilde{g}}$ is the same as the law of Z under $P_{2b}^{g-\tilde{g}}$. □

We extend the Markov process (Z, P_b^g) to the state space $G_\dagger = G \cup \{\dagger\}$ by declaring that $\dagger$ is an absorbing point.

3. Coalescing processes. The purpose of this section is to introduce the coalescing Lévy process ϑ, a nonlocally coalescing Lévy process ${}^k\vartheta$ and the coalescing random walk ${}^k\overline{\vartheta}$, the instantaneously coalescing Lévy process η, and to relate these processes.

3.1. *Coalescing Lévy processes ϑ and ${}^k\vartheta$.* We will give a sample path construction of ϑ. In fact, we will couple the construction of ϑ with that of a sequence of *nonlocally* coalescing Lévy processes ${}^k\vartheta$, in which particles die at a rate proportional to the weighted amount of time they have spent within distance N^{-k} of other living particles.

Fix $n \in \mathbb{N}$ and $\mathbf{g} = (g_1, \dots, g_n) \in G_\dagger^n$. On some probability space with probability measure denoted by $\mathbf{P}^{\mathbf{g}} = \mathbf{P}_b^{\mathbf{g}}$, let $\mathbf{Z} = (Z_1, \dots, Z_n)$ be a vector of independent Lévy processes (with Lévy constant b) starting at $\mathbf{g}$. For $k \in \mathbb{Z}$, $1 \le i < j \le n$, with both g_i and g_j different from $\dagger$ and $t \ge 0$, we introduce the following *approximate collision local time* of Z_i and Z_j:

$${}^kL_{i,j}(t) := N^k \int_0^t ds\, \mathbf{1}\{|Z_i(s) - Z_j(s)| \le N^{-k}\}. \tag{36}$$

Note that the limit

$${}^\infty L_{i,j}(t) := \lim_{k\to\infty} {}^kL_{i,j}(t) \quad \text{uniformly on compacts, } \mathbf{P}_b^{\mathbf{g}}\text{-a.s.,} \tag{37}$$

is the *collision local time* of the ith and jth particles (Corollary 14). For the other pairs $1 \le i < j \le n$ such that $g_i = \dagger$ or $g_j = \dagger$ set ${}^kL_{i,j} \equiv 0$.

On the same probability space where $\mathbf{Z}$ is defined, suppose that we also have defined a family $\tau_{i,j}$, $1 \le i < j \le n$, of random variables that are *exponentially* distributed with mean 1, independent and jointly independent of $\mathbf{Z}$.

Recall that $a > 0$ is a given (diffusion) constant. For $k \in \overline{\mathbb{Z}} := \mathbb{Z} \cup \{\infty\}$ and $1 \le i < j \le n$, set

$$ {}^kU_{i,j} := \inf\{t\colon\ a\,({}^kL_{i,j}(t)) > \tau_{i,j}\}. \tag{38} $$

We will say that the jth particle *coalesces into* the ith one at time ${}^kU_{i,j}$, provided that, at time ${}^kU_{i,j}-$, both are still alive. That is, $a\,{}^kL_{i,j}$ serves as a clock under which the ith particle tries to kill the jth one, as long as both are not yet killed. To be more precise, recall that $\mathbf{Z}(0) = \mathbf{g} \in G_\dagger^n$ and define a $\{0,1\}^n$-valued càdlàg process ${}^k\mathbf{I} := \{({}^kI_j(t))_{1\le j\le n}\colon t \ge 0\}$ starting at

$$ {}^kI_j(0) := \begin{cases} 0, \text{ if } g_j \ne \dagger, \\ 1, \text{ if } g_j = \dagger, \end{cases} \qquad k \in \overline{\mathbb{Z}},\ 1 \le i \le n, $$

by setting

$$ {}^kI_j(t) := {}^kI_j(0) + \sum_{i<j} \mathbf{1}\{{}^kU_{i,j} \le t\}\big(1 - {}^kI_i({}^kU_{i,j}-)\big)\big(1 - {}^kI_j({}^kU_{i,j}-)\big). $$

As the ${}^kU_{i,j}$ are $\mathbf{P}^{\mathbf{g}}$-a.s. distinct, there is no problem with this definition. The interpretation is that ${}^kI_i(t)$ is the indicator of the event that at time t the ith particle is dead. Here we are allowing some particles to be already dead at time 0.

Define a $G_\dagger^n$-valued process ${}^k\vartheta := \{({}^k\vartheta_i(t))_{1\le i\le n}\colon t \ge 0\}$ starting at

$$ {}^k\vartheta_i(t) := \begin{cases} Z_i(t), & \text{if } {}^kI_i(t) = 0, \\ \dagger, & \text{if } {}^kI_i(t) = 1. \end{cases} \tag{39} $$

That is, killed particles are sent to $\dagger$ where they stay forever. Let ${}^k\mathbf{P}^{\mathbf{g}}_{a,b}$ denote the law of ${}^k\vartheta$ starting at $\mathbf{g} \in G_\dagger^n$. For $k \in \mathbb{Z}$, we call ${}^k\vartheta$ a *nonlocally coalescing Lévy process*, and drop the word "nonlocally" in the case of ${}^\infty\vartheta$. We also write simply $(\vartheta, \mathbf{P}^{\mathbf{g}}_{a,b})$ instead of $({}^\infty\vartheta, {}^\infty\mathbf{P}^{\mathbf{g}}_{a,b})$. The following result is immediate from the properties of $\mathbf{Z}$ and $\tau_{i,j}$.

LEMMA 15 [(Nonlocally) coalescing Lévy process]. *We have that* $({}^k\vartheta, {}^k\mathbf{P}^{\mathbf{g}}_{a,b})$ *is a time-homogeneous strong Markov process for each* $k \in \overline{\mathbb{Z}}$.

REMARK 16 (Feller property). Because we can write Z_i starting at $g_i \in G$ as $g_i + \widetilde{Z}_i$ with $\widetilde{Z}_i(0) = 0$ and because

$$ \lim_{k\to\infty} N^k \int_0^t ds\, \mathbf{1}\{|g_i + \widetilde{Z}_i(s) - (g_j + \widetilde{Z}_j(s))| \le N^{-k}\} $$

is continuous in $(t, g_i, g_j) \in \mathbb{R}_+ \times G^2$, by Proposition 13 it is not hard to demonstrate that ${}^k\vartheta$ is actually *Feller* for each $k \in \overline{\mathbb{Z}}$.

3.2. *Instantaneously coalescing Lévy processes* η. Fix $n \in \mathbb{N}$ and $\mathbf{g} = (g_1, \dots, g_n) \in \check{G}^n_\dagger$ (here $g_i = g_j \neq \dagger$ does not hold by definition; see the beginning of Section 1.5). Let $\mathbf{P}^{\mathbf{g}}$ and $\mathbf{Z}$ be as in Section 3.1.

For $1 \le i < j \le n$ with both g_i and g_j different from †, set

$$V_{i,j} = \inf\{t \ge 0 : Z_i(t) = Z_j(t)\}$$

for the hitting time of Z_i and Z_j. Recall that $V_{i,j} < \infty$ with $\mathbf{P}^{\mathbf{g}}$-probability 1 (Corollary 14). If $i' \notin \{i, j\}$, then $Z_{i'}(V_{i,j}) \neq Z_i(V_{i,j}) = Z_j(V_{i,j})$, $\mathbf{P}^{\mathbf{g}}$-a.s., by the independence of the coordinates of $\mathbf{Z}$ and the fact that the distribution of $Z_{i'}(t)$ is absolutely continuous for all $t > 0$ when $g_{i'} \neq \dagger$. In particular, $V_{i,j} \neq V_{i',j'}$, $\mathbf{P}^{\mathbf{g}}$-a.s., when $(i, j) \neq (i', j')$. For $1 \le i < j \le n$ such that $g_i = \dagger$ or $g_j = \dagger$, put $V_{i,j} := \infty$.

Define a $\{0,1\}^n$-valued càdlàg process $\mathbf{J} := \{(J_j(t))_{1 \le j \le n} : t \ge 0\}$ starting at

$$J_j(0) := \begin{cases} 0, & \text{if } g_j \neq \dagger, \\ 1, & \text{if } g_j = \dagger, \end{cases} \qquad 1 \le j \le n,$$

by setting

$$J_j(t) := J_j(0) + \sum_{i<j} \mathbf{1}\{V_{i,j} \le t\}\big(1 - J_i(V_{i,j}-)\big)\big(1 - J_j(V_{i,j}-)\big).$$

As the $V_{i,j}$ are $\mathbf{P}^{\mathbf{g}}$-a.s. distinct, there is again no problem with this definition, and the interpretation is that $J_i(t)$ is the indicator of the event that at time t the ith particle is dead. Here we are allowing some particles to be already dead at time 0.

Define a $\check{G}^n_\dagger$-valued process $\eta := \{(\eta_i(t))_{1 \le i \le n} : t \ge 0\}$ starting at $\mathbf{g} \in \check{G}^n_\dagger$ by

$$\eta_i(t) := \begin{cases} Z_i(t), & \text{if } J_i(t) = 0, \\ \dagger, & \text{if } J_i(t) = 1, \end{cases} \tag{40}$$

and denote its law by $\mathbf{Q}^{\mathbf{g}} = \mathbf{Q}^{\mathbf{g}}_b$. We call $(\eta, \mathbf{Q}^{\mathbf{g}}_b)$ an *instantaneously coalescing Lévy process*. The following result is immediate by construction.

LEMMA 17 (Instantaneously coalescing Lévy process). *We have that* $(\eta, \mathbf{Q}^{\mathbf{g}}_b)$ *is a time-homogeneous strong Markov process.*

3.3. *An absolute continuity property of* ϑ *and* η. Consider the coalescing Lévy process ϑ with $\vartheta(0) \neq \dagger$ [i.e., at least one of the $\vartheta_i(0)$ is different from †]. Let $R_t \subset \mathbb{N}$ denote the set of all labels of particles alive at time t; that is, $R_t := \{i : \vartheta_i(t) \neq \dagger\}$. Write $|R_t|$ for its cardinality. Define S_t analogously for the instantaneously coalescing Lévy process η.

LEMMA 18 (Absolute continuity). *Let $n \in \mathbb{N}$ and $\mathbf{g} \in G_\dagger^n$ with $\mathbf{g} \neq \dagger$. Take $\varnothing \neq R \subseteq \{1, \dots, n\}$ and $t > 0$. Then the (subprobability) measure*

$$\mathbf{P}^{\mathbf{g}}_{a,b}\big(R_t = R,\ (\vartheta_i(t))_{i \in R} \in d\mathbf{h}\big)$$

on $G^{|R|}$ is absolutely continuous with respect to the Haar measure $l^{|R|}$ on $G^{|R|}$ and, in fact, has a (subprobability) density function $\tilde{p}_t(\mathbf{g}, R\,;\,\cdot)$ that satisfies

$$\tilde{p}_t(\mathbf{g}, R\,;\,\mathbf{h}) \le \prod_{i \in R} p_t(h_i - g_i), \qquad \mathbf{h} \in G^{|R|}$$

(with p the transition density of the underlying Lévy process). An analogous result holds for S_t with the resulting density function being denoted by $\tilde{q}_t(\mathbf{g}, S; \cdot)$.

PROOF. For a Borel subset B of $G^{|R|}$,

$$\mathbf{P}^{\mathbf{g}}_{a,b}\big(R_t = R,\ (\vartheta_i(t))_{i \in R} \in B\big) \le \mathbf{P}^{\mathbf{g}}_{b}\big((Z_i(t))_{i \in R} \in B\big).$$

This implies the claim. □

3.4. *Coalescing random walk ${}^k\overline{\vartheta}$.* For each $k \in \mathbb{Z}$, the quotient map π_k from G to G/G_k transforms the Lévy process Z on G to a *random walk* ${}^k\bar{Z} := \pi_k Z$ on G/G_k. In order to calculate the jump rates of ${}^k\bar{Z}$, recall that the Haar measure l assigns mass N^{-k} to G_k and each of its cosets [see (6)]. Furthermore, if g belongs to a coset of G_k other than G_k itself, then $|g| = |\bar{g}|$, where $\bar{g} = \pi_k g \in G/G_k$ [recall (16)]. Hence, by the definition (10) of ν, the jump $\bar{g} \neq 0$ occurs in the walk ${}^k\bar{Z}$ with rate

$$(41)\qquad {}^k q_{\bar{g}} := bN^{-k}|\bar{g}|^{-\alpha-1}, \qquad \bar{g} \in G/G_k,\ \bar{g} \neq 0.$$

Note that the total jump rate is finite: $\sum_{\bar{g} \neq 0} {}^k q_{\bar{g}} = \nu(G \backslash G_k) < \infty$ [recall (26)].

If in the construction of Section 3.1 we put ${}^k\bar{\mathbf{Z}} := (\pi_k Z_1, \dots, \pi_k Z_n)$, then for pairs (i, j), $i < j$, such that both $g_i \neq \dagger$ and $g_j \neq \dagger$, by (36) we have

$$(42)\qquad {}^k L_{i,j}(t) = N^k \int_0^t ds\, \mathbf{1}\{{}^k\bar{Z}_i(s) = {}^k\bar{Z}_j(s)\}.$$

That is, ${}^k L_{i,j}$ from (36) is now the "weighted" collision local time of ${}^k\bar{Z}_i$ and ${}^k\bar{Z}_j$.

Recall (18) saying that G/G_k is isomorphic to the countable hierarchical group Ξ. Delayed coalescing random walks on Ξ are described in [12] and [10] as systems of unlabeled particles. As we remarked in Section 1.4 for the case of the usual description of (delayed) coalescing Lévy processes, it is possible to enhance such a model by assigning labels to the particles and, rather than thinking of two particles merging into one, think instead of one of the particles being sent to the cemetery † at the time of "coalescence." It is this latter process that we will refer to as a (*delayed*) *coalescing random walk* ${}^k\overline{\vartheta}$ on $G/G_k \cup \{\dagger\}$.

Combining the above observations and taking into account in particular the identity (42) leads immediately to the following result.

LEMMA 19 (Coalescing random walk). *Let $n \in \mathbb{N}$, $\mathbf{g} \in G_{\dagger}^{n}$ and $k \in \mathbb{Z}$. Under ${}^{k}\mathbf{P}_{a,b}^{\mathbf{g}}$, the process ${}^{k}\overline{\vartheta} := (\pi_k {}^{k}\vartheta_1, \dots, \pi_k {}^{k}\vartheta_n)$ is a coalescing random walk on $G/G_k \cup \{\dagger\}$ with jump rates ${}^{k}q$ of* (41), *coalescing rate aN^k and initial state $\bar{\mathbf{g}} = (\bar{g}_1, \dots, \bar{g}_n) = (\pi_k g_1, \dots, \pi_k g_n)$.*

3.5. *Convergence of coalescing processes.* In this section we will make precise one sense in which the coalescing random walks ${}^{k}\overline{\vartheta}$ converge to the coalescing Lévy process ϑ as $k \to \infty$, resp., the coalescing Lévy process ϑ tends to the instantaneously coalescing Lévy process η as $a \to \infty$.

Recall the definition of the state space $\mathbf{B}$ given in Section 1.4. For $k \in \mathbb{Z}$, define the *averaging transformation* $M_k : \mathbf{B} \to \mathbf{B}$ by

$$(M_k x)(g) := N^k \int_{G_k} dg'\, x(g' + g), \qquad g \in G. \tag{43}$$

That is, $M_k x(g)$ is the average of x over the coset $g + G_k$. Note that M_k is well defined as a map from $\mathbf{B}$ into itself because the right-hand side of (43) does not depend on which particular representative for x we use to compute the integral. Since $M_k x$ is constant on the cosets of G_k, we can think of $M_k x(\cdot)$ as a function on the quotient group G/G_k and write $\bar{M}_k x$ instead of $M_k x$ in this case.

By analogy with the product brackets pairing of Definition 1, we can introduce a pairing between $[0,1]^{G/G_k}$ and $(G/G_k \cup \{\dagger\})^n$, $n \in \mathbb{N}$, that we will also denote by $[\cdot,\ \cdot]$.

Recalling Lemma 19, the convergence of the coalescing random walk ${}^{k}\overline{\vartheta}$ to that of the coalescing Lévy process ϑ and the convergence of ϑ to the instantaneously coalescing Lévy process η can now be expressed as follows.

PROPOSITION 20 (Convergence). *Suppose $n \in \mathbb{N}$, $\varphi \in L^1(G^n, l^n)$ and $t \geq 0$. Then*

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, {}^{k}\mathbf{P}_{a,b}^{\mathbf{g}}[\bar{M}_k x,\ {}^{k}\overline{\vartheta}_t] \to \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, \mathbf{P}_{a,b}^{\mathbf{g}}[x, \vartheta_t] \quad \text{as } k \to \infty \tag{44}$$

and

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{P}_{a,b}^{\mathbf{g}}[x, \vartheta_t] \to \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{Q}_{b}^{\mathbf{g}}[x, \eta_t] \quad \text{as } a \to \infty, \tag{45}$$

uniformly in $x \in \mathbf{B}$.

PROOF. Fix n, φ and t as in the proposition. First consider (44). Note that the right-hand side of (44) is well defined (i.e., does not depend on which particular representative we choose for x) by Lemma 18. Using the definition (43) of the average $M_k x$, the construction of ${}^{k}\vartheta_t$ provided in Section 3.1 and interchanging the order of expectation and integration, the left-hand side of (44) can be written as

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, N^{nk} \int_{G_k^n} d\mathbf{h}\, \mathbf{P}_b^{\mathbf{g}} \prod_{i=1}^{n} x^{1 - {}^{k}I_i(t)}(Z_i(t) - h_i).$$

Since for $\mathbf{h} \in G_k^n$ the law of $(\mathbf{Z} - \mathbf{h}, {}^k\mathbf{I})$ under $\mathbf{P}_b^{\mathbf{g}}$ is the same as the law of $(\mathbf{Z}, {}^k\mathbf{I})$ under $\mathbf{P}_b^{\mathbf{g}-\mathbf{h}}$, the latter expectation equals

$$\mathbf{P}_b^{\mathbf{g}-\mathbf{h}} \prod_{i=1}^n x^{1-{}^kI_i(t)}(Z_i(t)).$$

Interchanging the order of integration (twice) and using the shift invariance of the Haar measure $d\mathbf{g}$, the left-hand side of (44) can be rewritten as

$$\int_{G^n} d\mathbf{g}\, N^{nk} \int_{G_k^n} d\mathbf{h}\, \varphi(\mathbf{g}+\mathbf{h}) \mathbf{P}_b^{\mathbf{g}} \prod_{i=1}^n x^{1-{}^kI_i(t)}(Z_i(t)).$$

The difference between the left-hand side and the right-hand side of (44) can be written as a sum of two terms by subtracting and adding the quantity

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{P}_b^{\mathbf{g}} \prod_{i=1}^n x^{1-{}^kI_i(t)}(Z_i(t)). \tag{46}$$

The absolute value of the first term in this sum can be estimated from above by

$$\int_{G^n} d\mathbf{g} \left| N^{nk} \int_{G_k^n} d\mathbf{h}\, \varphi(\mathbf{g}+\mathbf{h}) - \varphi(\mathbf{g}) \right| \to 0 \quad \text{as } k \to \infty,$$

where the convergence follows from Lemma 8. It therefore remains to check that (46) converges uniformly in $x \in \mathbf{B}$ to

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{P}_b^{\mathbf{g}} \prod_{i=1}^n x^{1-{}^\infty I_i(t)}(Z_i(t))$$

as $k \to \infty$. Note that our fixed $t \geq 0$ is $\mathbf{P}_b^{\mathbf{g}}$-a.s. different from ${}^\infty U_{i,j}$ [recall (38)] for any $1 \leq i < j \leq n$, and these random variables are $\mathbf{P}_b^{\mathbf{g}}$-a.s. distinct. Moreover, ${}^kL_{i,j}(t)$ converges uniformly on compacts to ${}^\infty L_{i,j}(t)$ as $k \to \infty$, $\mathbf{P}_b^{\mathbf{g}}$-a.s. [recall (37)]. Thus, the ${}^kU_{i,j}$ converge $\mathbf{P}_b^{\mathbf{g}}$-a.s. to the ${}^\infty U_{i,j}$ as $k \to \infty$, and we have $\mathbf{P}_b^{\mathbf{g}}$-a.s. for $1 \leq i \leq n$ that ${}^kI_i(t) = {}^\infty I_i(t)$ for all $k \in \mathbb{Z}$ sufficiently large.

The proof of (45) is similar and easier. Write ${}^\infty_a U_{i,j}$ and ${}^\infty_a I_i$ in place of ${}^\infty U_{i,j}$ and ${}^\infty I_i$ for the moment, to emphasize the dependence on a in the definition. We need to check that

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{P}_b^{\mathbf{g}} \prod_{i=1}^n x^{1-{}^\infty_a I_i(t)}(Z_i(t))$$

converges uniformly to

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{Q}_b^{\mathbf{g}} \prod_{i=1}^n x^{1-J_i(t)}(Z_i(t))$$

as $a \to \infty$. It follows from Corollary 14 that $\mathbf{P}_b^{\mathbf{g}}$-a.s. the random variable ${}^\infty_a U_{i,j}$ converges to the hitting time $V_{i,j}$ as $a \to \infty$. An argument similar to the one above establishes we have $\mathbf{P}_b^{\mathbf{g}}$-a.s. for $1 \le i \le n$ that ${}^{\infty}_{a}I_i(t) = J_i(t)$ for all a sufficiently large, implying the claim. □

3.6. *Scaling of coalescing processes.* The shrinking automorphism σ defined in the beginning of Section 1.6 can be extended to $G_\dagger$ by setting $\sigma\dagger := \dagger$, and to $G_\dagger^n$ by $\sigma(g_1, \dots, g_n) := (\sigma g_1, \dots, \sigma g_n)$ for each $n \in \mathbb{N}$.

LEMMA 21 (Scaling for ϑ and η). *For $m \in \mathbb{Z}$, $s \in \mathbb{R}$ and $\mathbf{g} \in G_\dagger^n$, $n \in \mathbb{N}$, the distribution of the process $\vartheta(N^{\alpha s}\cdot)$ under $\mathbf{P}_{a,b}^{\sigma^{-m}\mathbf{g}}$ [resp. $\eta(N^{\alpha s}\cdot)$ under $\mathbf{Q}_b^{\sigma^{-m}\mathbf{g}}$] is the same as the distribution of the process $\sigma^{-m}\vartheta$ under $\mathbf{P}^{\mathbf{g}}_{N^{\alpha s-m}a,\, N^{\alpha(s-m)}b}$ (resp. $\sigma^{-m}\eta$ under $\mathbf{Q}^{\mathbf{g}}_{N^{\alpha(s-m)}b}$).*

PROOF. We will consider the claim for ϑ. The proof for η is similar and is omitted. In the notation of Section 3.1, we have from Corollary 10 that the distribution of $\mathbf{Z}(N^{\alpha s}\cdot)$ under $\mathbf{P}_b^{\sigma^{-m}\mathbf{g}}$ is the same as the distribution of $\sigma^{-m}\mathbf{Z}$ under $\mathbf{P}^{\mathbf{g}}_{N^{\alpha(s-m)}b}$. Therefore, the distribution of $(\mathbf{Z}(N^{\alpha s}\cdot),\ ({}^{\infty}L_{i,j}(N^{\alpha s}\cdot))_{1\le i<j\le n})$ under $\mathbf{P}_b^{\sigma^{-m}\mathbf{g}}$ is the same as the distribution of $(\sigma^{-m}\mathbf{Z},\ (N^{\alpha s-m}\,{}^{\infty}L_{i,j})_{1\le i<j\le n})$ under $\mathbf{P}^{\mathbf{g}}_{N^{\alpha(s-m)}}$, and the result is immediate from the construction of Section 3.1. □

4. Existence and uniqueness for *X* and *Y*. This section is devoted to the proof of Theorems 3 and 4. We begin with the following simple observation. Recall the function I_n^φ of (11).

LEMMA 22 (The algebra $\mathbf{A}$). *Let $\mathscr{A} \subset C_c(G)$ denote the set of functions of the form $\mathbf{1}_H$, where H is a coset of G_k for some $k \in \mathbb{Z}$, and write $\mathbf{A}$ for the linear span of the set*

$$\left\{I_n^\varphi : n \in \mathbb{N}, \varphi = \bigotimes_{i=1}^n \varphi_i, \varphi_i \in \mathscr{A}\right\} \cup \{\textit{constant functions on } \mathbf{B}\}.$$

Then $\mathbf{A}$ is a dense subspace of $C(\mathbf{B})$.

PROOF. The result will be immediate from the Stone–Weierstrass theorem (see, e.g., Theorem 36A of Simmons [20]), once we know that the algebra $\mathbf{A}$ separates points. However, if for $x_1, x_2 \in \mathbf{B}$,

$$\int_G dg\, x_1(g)\,\varphi(g) = \int_G dg\, x_2(g)\varphi(g), \qquad \varphi \in \mathscr{A},$$

then $x_1 = x_2$. □

PROOF OF THEOREM 3.

Step 1. *Reformulation of the right-hand side of* (12). Fix t, n, φ, x as in the theorem. Recall the notation R_t (introduced in Section 3.3) for the set of all labels of particles of ϑ alive at time t. Decompose the right-hand side of (12) into a sum with $2^n - 1$ terms by introducing into the expectation expression under the integral the indicator functions $\mathbf{1}\{R_t = R\}$ for $\varnothing \neq R \subseteq \{1, \dots, n\}$. By Lemma 18 we know that ϑ_t restricted to $\{R_t = R\}$ has an absolutely continuous subprobability distribution with density function $\tilde{p}_t(\mathbf{g}, R; \cdot)$. Hence, for a typical summand we get

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{P}^{\mathbf{g}} \mathbf{1}\{R_t = R\}[x, \vartheta_t] = \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \int_{G^{|R|}} d\mathbf{h}\; \tilde{p}_t(\mathbf{g}, R\,; \mathbf{h})\,[x, \mathbf{h}].$$

Introduce the function

$$\varphi_t^R(\mathbf{h}) := \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, \tilde{p}_t(\mathbf{g}, R;, \mathbf{h}), \qquad \mathbf{h} \in G^{|R|}. \tag{47}$$

Note that it belongs to $L^1(G^{|R|}, l^{|R|})$. In fact, since the $\tilde{p}_t(\mathbf{g}, R\,; \cdot\,)$ are subprobability densities,

$$\int_{G^{|R|}} d\mathbf{h} \big|\varphi_t^R(\mathbf{h})\big| \le \int_{G^{|R|}} d\mathbf{h} \int_{G^n} d\mathbf{g}\; |\varphi(\mathbf{g})|\, \tilde{p}_t(\mathbf{g}, R; \mathbf{h}) \le \int_{G^n} d\mathbf{g}\; |\varphi(\mathbf{g})| < \infty.$$

Using this function, the right-hand side of (12) can thus be written as

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \mathbf{P}^{\mathbf{g}}[x, \vartheta_t] = \sum_R I_{|R|}^{\varphi_t^R}(x) \in C(\mathbf{B}). \tag{48}$$

In particular, we see that the right-hand side of (12) is well defined (i.e., it does not depend on the choice of the representative of x).

Step 2. *Uniqueness*. By Lemma 22 we know that at most one semigroup exists with the required properties.

Step 3. *Existence of transition kernels*. Fix $k \in \mathbb{Z}$. Using the isomorphism $G/G_k \cong \Xi$ [recall (18)], we may make use of the well-known model of *interacting Fisher–Wright diffusions* labeled by the countable hierarchical group Ξ (see, e.g., [12] or [10]). Define ${}^k\overline{X}$ to be such a process with the *resampling mechanism* given by $N^k af$, where f is as given by (9), and with *migration* determined by the random walk ${}^k\bar{Z}$ in G/G_k introduced in the beginning of Section 3.4.

More precisely, given the starting point $\bar{x} \in [0,1]^{G/G_k}$, we may construct ${}^k\overline{X}$ as the unique strong solution of the following $[0,1]^{G/G_k}$-valued system of stochastic differential equations:

$$\begin{aligned} {}^k\overline{X}_0(\bar{g}) &= \bar{x}(\bar{g}), \\ d\,{}^k\overline{X}_t(\bar{g}) &= \sqrt{N^k af\big({}^k\overline{X}_t(\bar{g})\big)}\, w(dt, \bar{g}) \\ &\qquad + \sum_{\bar{g}' \neq 0} {}^k q_{\bar{g}'} \big[{}^k\overline{X}_t(\bar{g} + \bar{g}') - {}^k\overline{X}_t(\bar{g})\big]\, dt, \end{aligned}$$

$\bar{g} \in G/G_k$, where $w(\cdot, \bar{g})$, $\bar{g} \in G/G_k$, are i.i.d. standard Brownian motions and the migration rates ${}^k q_{\bar{g}'}$ are given by (41). Write ${}^k\mathbb{P}_{a,b}^{\bar{x}}$ for the law of ${}^k\overline{X}$ starting at $\bar{x} \in [0,1]^{G/G_k}$.

Shiga's [18] *duality relation* between interacting Fisher–Wright diffusions and coalescing random walks may be expressed in our notation as follows. For $k \in \mathbb{Z}$, $n \in \mathbb{N}$, $t \geq 0$, $\bar{x} \in [0,1]^{G/G_k}$ and $\bar{\mathbf{g}} \in (G/G_k)^n$ of the form $\bar{\mathbf{g}} = (\bar{g}_1, \dots, \bar{g}_n) = (\pi_k g_1, \dots, \pi_k g_n) = \pi_k \mathbf{g}$ for $\mathbf{g} = (g_1, \dots, g_n) \in G$,

$$ {}^k\mathbb{P}^{\bar{x}}_{a,b}\big[{}^k\overline{X}_t, \bar{\mathbf{g}} \big] = {}^k\mathbf{P}^{\mathbf{g}}_{a,b}\big[\bar{x}, {}^k\overline{\vartheta}_t\big]. \tag{49} $$

Recall Proposition 20. Using the duality observation (49), we may rewrite the left-hand side in the convergence statement (44) as

$$ \int_{G^n} d\mathbf{g}\; \varphi(\mathbf{g})\, {}^k\mathbb{P}^{\bar{M}_k x}_{a,b}\big[{}^k\overline{X}_t,\ \pi_k \mathbf{g} \big]. \tag{50} $$

In order to express this in terms of the functions I^{φ}_n from (11), we introduce the *liftings* $L_k\colon [0,1]^{G/G_k} \to \mathbf{B}$ defined by

$$ (L_k \bar{x})(g) := \bar{x}(\pi_k g), \qquad \bar{x} \in [0,\, 1]^{G/G_k}, \qquad g \in G. $$

Observe that the composition $\bar{M}_k \circ L_k$ is the identity map on $[0,\, 1]^{G/G_k}$, whereas $L_k \circ \bar{M}_k = M_k$ on G. Now (50) and hence the left-hand side of (44) equals

$$ {}^k\mathbb{P}^{\bar{M}_k x}_{a,b} I^{\varphi}_n\big(L_k\, {}^k\overline{X}_t \big) =: \int {}^k\nu_t(x, dy) I^{\varphi}_n(y), $$

where ${}^k\nu_t(x, \cdot)$ denotes the distribution of $L^k_k\overline{X}_t$ under ${}^k\mathbb{P}^{\bar{M}_k x}_{a,b}$. From the convergence statement (44), linearity and Lemma 22, we conclude that there exist probability laws $\nu_t(x, \cdot)$ on $\mathbf{B}$, such that

$$ \int \nu_t(x, dy) I^{\varphi}_n(y) = \int_{G^n} d\mathbf{g}\; \varphi(\mathbf{g}) \mathbf{P}^{\mathbf{g}}_{a,b}[x, \vartheta_t]. \tag{51} $$

Step 4. *Feller property*. It is immediate from (48) and (51) that, for $t \geq 0$, $n \in \mathbb{N}$ and $\varphi \in L^1(G^n, l^n)$, the map $\mathbf{B} \ni x \mapsto \int \nu_t(x, dy)\, I^{\varphi}_n(y)$ is continuous. Therefore, by linearity and Lemma 22, there is an operator $\mathsf{S}_{a,b}(t)\colon C(\mathbf{B}) \to C(\mathbf{B})$ such that

$$ \mathbf{B} \ni x \mapsto \int \nu_t(x, dy)\, F(y) = \mathsf{S}_{a,b}(t) F(x) \tag{52} $$

for $F \in C(\mathbf{B})$, and $\mathsf{S}_{a,b}(t)$ satisfies (12).

Step 5. *Semigroup property*. Now we want to check the Chapman–Kolmogorov property of the kernels $\nu_t(x, dy)$ from (51). It suffices to show that

$$ \int \nu_t(x, dy) \int \nu_s(y, dz)\, I^{\varphi}_n(z) = \int \nu_{t+s}(x, dz)\, I^{\varphi}_n(z). \tag{53} $$

According to (51) and (48), the interior integral can be rewritten to get, for the left-hand side of (53),

$$ \sum_R \int \nu_t(x, dy)\, I^{\varphi^R_s}_{|R|}(y). $$

Again by (51) we may continue with

$$\sum_R \int_{G^{|R|}} d\mathbf{h}\ \varphi_s^R(\mathbf{h})\mathbf{P}^{\mathbf{h}}_{a,b}[x, \vartheta_t].$$

Inserting (47) and interchanging the order of integration leads to

$$\int_{G^n} d\mathbf{g}\ \varphi(\mathbf{g}) \sum_R \int_{G^{|R|}} d\mathbf{h}\ \tilde{p}_s(\mathbf{g}, R; \mathbf{h})\mathbf{P}^{\mathbf{h}}_{a,b}[x, \vartheta_t].$$

Applying the Markov property of ϑ and (51) once more, we arrive at the right-hand side of (53).

Step 6. *Strong continuity*. We have established the existence of a Markov semigroup of operators $\mathsf{S}_{a,b}(t): C(\mathbf{B}) \to C(\mathbf{B})$. In order to show that this semigroup $\mathsf{S}_{a,b}$ is strongly continuous, it suffices, by the Remark after Theorem 1.9.4 of Blumenthal and Getoor [2], to show that

$$\lim_{t\downarrow 0} \mathsf{S}_{a,b}(t)F(x) = F(x), \qquad F \in C(\mathbf{B}),\ x \in \mathbf{B}. \tag{54}$$

By linearity and Lemma 22, it in turn suffices to check (54) for $F = I_n^{\varphi}$ for all $\varphi \in C(G^n)$, $n \in \mathbb{N}$. Write a typical term from the left-hand side of (54) as in (12). Recalling the construction of ϑ in Section 3.1, we will use

$$\begin{aligned} &\left| \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\, \mathbf{P}^{\mathbf{g}}_{a,b}[x, \vartheta_t] - \int_{G^n} d\mathbf{g}\ \varphi(\mathbf{g})\, \mathbf{P}^{\mathbf{g}}[x, \mathbf{Z}(t)] \right| \\ &\qquad \le \text{const} \int_{G^n} d\mathbf{g}\ \varphi(\mathbf{g})\, \mathbf{P}^{\mathbf{g}}\{\text{a coalescence has occurred by time } t\}. \end{aligned} \tag{55}$$

Since this tends to 0 as $t \downarrow 0$, it suffices to replace ϑ by $\mathbf{Z}$, that is, to consider the second term in (55). Reversing time, we obtain

$$\int_{G^n} d\mathbf{h}\ [x, \mathbf{h}]\, \mathbf{P}^{\mathbf{h}} \varphi(\mathbf{Z}(t)) \to \int_{G^n} d\mathbf{h}[x, \mathbf{h}]\, \varphi(\mathbf{h}) = I_n^{\varphi}(x) \quad \text{as } t \downarrow 0,$$

as required.

Step 7. *Hunt process*. From general Markov theory (see, e.g., Theorem 1.9.4 of Blumenthal and Getoor [2]), we can conclude from steps 1–6 that there is a Hunt process $(X, \mathbb{P}^x_{a,b})$ with semigroup $\mathsf{S}_{a,b}$.

Step 8. *Continuous sample paths*. The general theory only yields that the Hunt process X has càdlàg paths. In order to show that X has continuous paths, it will suffice to show that the distribution of the continuous process $L_k\, {}^k\overline{X}$ under ${}^k\mathbb{P}^{\bar{M}_k x}_{a,b}$ converges to the distribution of X under $\mathbb{P}^x_{a,b}$ in the sense of convergence of distributions on the Skorohod space $D(\mathbb{R}_+,\ \mathbf{B})$ (cf. Theorem 3.10.2 of Ethier and Kurtz [7]).

Arguments similar to those in Steps 1–6 establish that $\left({}^k\overline{X},\ {}^k\mathbb{P}^{\bar{x}}_{a,b}\right)$ has a Feller semigroup. The latter convergence statement will follow from Theorem

4.2.11 of [7] if we can show that

$$\sup_{\bar{x}\in[0,1]^{G/G_k}} \left| {}^k\mathbb{P}^{\bar{x}}_{a,b} I^{\varphi}_n\big(L_k\, {}^k\overline{X}_t\big) - \mathbb{P}^{L_k\bar{x}}_{a,b} I^{\varphi}_n(X_t)\right| \to 0 \quad \text{as } k\to\infty. \tag{56}$$

The supremum can also be written as

$$\sup_{x\in\mathbf{B}} \left| {}^k\mathbb{P}^{\bar{M}_k x}_{a,b} I^{\varphi}_n\big(L_k\, {}^k\overline{X}_t\big) - \mathbb{P}^{M_k x}_{a,b} I^{\varphi}_n(X_t)\right|.$$

It follows from Proposition 20 that

$$\sup_{x\in\mathbf{B}} \left| {}^k\mathbb{P}^{\bar{M}_k x}_{a,b} I^{\varphi}_n\big(L_k\, {}^k\overline{X}_t\big) - \mathbb{P}^{x}_{a,b} I^{\varphi}_n(X_t)\right| \to 0 \quad \text{as } k\to\infty.$$

Using Lemma 18, we have

$$\mathbb{P}^x_{a,b} I^{\varphi}_n(X_t) = \sum_R \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \int_{G^{|R|}} d\mathbf{h}\, \tilde{p}_t(\mathbf{g}, R; \mathbf{h})\,[x, \mathbf{h}]$$

and

$$\mathbb{P}^{M_k x}_{a,b} I^{\varphi}_n(X_t) = \sum_R \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g}) \int_{G^{|R|}} d\mathbf{h}\, \tilde{p}_t(\mathbf{g}, R; \mathbf{h}) N^{|R|k} \int_{G_k^{|R|}} d\mathbf{h}'\, [x, \mathbf{h}+\mathbf{h}'].$$

Thus, we can bound $\left|\mathbb{P}^{M_k x}_{a,b} I^{\varphi}_n(X_t) - \mathbb{P}^{x}_{a,b} I^{\varphi}_n(X_t)\right|$ above by

$$\sum_R \int_{G^n} d\mathbf{g} |\varphi(\mathbf{g})| \int_{G^{|R|}} d/\mathbf{h} \left| \tilde{p}_t(\mathbf{g}, R; \mathbf{h}) - N^{|R|k} \int_{G_k^{|R|}} d\mathbf{h}'\, \tilde{p}_t(\mathbf{g}, R; \mathbf{h}-\mathbf{h}')\right|.$$

By Lemma 8, the internal integral converges to 0 as $k\to\infty$ for l^n-almost all $\mathbf{g}\in G^n$. Therefore, by dominated convergence,

$$\sup_{x\in\mathbf{B}} \left|\mathbb{P}^{M_k x}_{a,b} I^{\varphi}_n(X_t) - \mathbb{P}^{x}_{a,b} I^{\varphi}_n(X_t)\right| \to 0 \quad \text{as } k\to\infty,$$

hence, (56) holds. □

PROOF OF THEOREM 4. The proof is very similar to that of Theorem 3 and rather easier, so we will omit the details. Essentially, we just replace the occurrences of (44) and ${}^k\overline{X}$ in the above proof by (45) and X, respectively. As X and Y have the same state space, there is no need for an analog of the liftings, L_k, and so in the counterpart of Step 8 it is possible to replace the application of Theorem 4.2.11 of [7] by one of Theorem 4.2.5 of [7]. □

5. Scaling results. The purpose of this section is to verify the cluster formation theorem (Theorem 6). This requires the following preparation.

PROOF OF PROPOSITION 5. Consider first the claim regarding X. Fix m, $s\in\mathbb{Z}$. A simple induction argument shows that it suffices to establish, for fixed $t>0$,

$$\mathsf{S}_{a,b}(N^{\alpha s}t)(F\circ\sigma^{-m}) = \big(\mathsf{S}_{N^{\alpha s-m}a,\, N^{\alpha(s-m)}b}(t)F\big)\circ\sigma^{-m}$$

for all $F \in C(B)$. By Lemma 22, it in turn suffices to consider the special case $F = I_n^\varphi$ for $n \in \mathbb{N}$ and $\varphi \in L^1(G^n, l^n)$.

Observe that, by definition of the shrinking operation,

$$\int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\,[x, \sigma^{-m}\mathbf{g}] = N^{mn} \int_{G^n} d\mathbf{g}\ \varphi(\sigma^m \mathbf{g})\,[x, \mathbf{g}], \qquad x \in \mathbf{B}.$$

Hence, by the definition (11) of I_n^φ, we get $I_n^\varphi \circ \sigma^{-m} = N^{mn} I_n^{\varphi \circ \sigma^m}(x)$. Thus, by (12),

$$\begin{aligned} \mathsf{S}_{a,b}(N^{\alpha s}t)(I_n^\varphi \circ \sigma^{-m})(x) &= N^{mn} \int_{G^n} d\mathbf{g}\ \varphi(\sigma^m \mathbf{g})\,\mathbf{P}^{\mathbf{g}}_{a,b}\big[x, \vartheta(N^{\alpha s}t)\big] \\ &= \int_{G^n} d\mathbf{g}\ \varphi(\mathbf{g})\,\mathbf{P}^{\sigma^{-m}\mathbf{g}}_{a,b}\big[x, \vartheta(N^{\alpha s}t)\big]. \end{aligned}$$

By Lemma 21, we may continue with

$$\begin{aligned} &= \int_{G^n} d\mathbf{g}\, \varphi(\mathbf{g})\,\mathbf{P}^{\mathbf{g}}_{N^{\alpha s - m}a,\, N^{\alpha(s-m)}b}[x, \sigma^{-m}\vartheta(t)] \\ &= (\mathsf{S}_{N^{\alpha s - m}a,\, N^{\alpha(s-m)}b} I_n^\varphi) \circ \sigma^{-m}. \quad \square \end{aligned}$$

LEMMA 23 (Ergodic theorem). *Let μ be as in Theorem 6. Then the sequence of probability measures $\sigma^{-m}\mu$ converges weakly to the point mass $\delta_{\theta \mathbf{1}}$ as $m \to \infty$.*

PROOF. For $k \in \mathbb{Z}$ and $m \geq k$, in $L^2 = L^2(\mathbf{B}, \mu)$ we have

$$\begin{aligned} \int_{G_k} dg\ \sigma^{-m}x(g) &= N^{-m} \int_{G_{k-m}} dg\ x(g) \\ &= N^{-k} N^{k-m} \int_{G_{k-m}} dg \int_{G_0} dh\ x(g+h) \end{aligned}$$

by the stationarity of μ. Since μ is ergodic, from the L^2-ergodic theorem (Theorem 6.4.1) of Krengel [13], it follows that the latter expression converges in L^2 to

$$N^{-k} \int \mu(dx) \int_{G_0} dh\, x(h) = \theta \int_{G_k} dh$$

as $m \to \infty$, where we used the assumption (15). Consequently, if H is a coset of G_k we have that

$$\int dg\, \mathbf{1}_H(g)\, \sigma^{-m}x(g) \to \theta \int dg\ \mathbf{1}_H(g) \quad \text{as } m \to \infty$$

in L^2. Thus, in the notation of Lemma 22 we get

$$\int \sigma^{-m}\mu(dx)\ I_n^\varphi(x) \to \int \delta_{\theta\mathbf{1}}(dx)\, I_n^\varphi(x) \quad \text{as } m \to \infty$$

for $\varphi = \bigotimes_{i=1}^n \varphi_i$ with $\varphi_i \in \mathscr{A}$, $1 \leq i \leq n$, $n \in \mathbb{N}$, and the result follows by Lemma 22. $\square$

PROOF OF THEOREM 6. (i) This follows, directly from Proposition 5, Lemma 23, Theorem 4 and Theorem 4.2.5 of [7].

(ii) This is immediate from part (i) and the observation that $\Theta_{-j,0}\circ\Theta_{m,m} = \Theta_{m-j,m}$.

(iii) From Proposition 5 we see that the distribution of $\Theta_{-j,0}Y$ under $\mathbb{Q}_b^{\theta\mathbf{1}}$ is the same as the distribution of Y under $\mathbb{Q}_{N^{\alpha j}b}^{\theta\mathbf{1}}$. For $\varphi \in L^1(G^n, l^n)$, $n \in \mathbb{N}$, by (13) we have

$$\mathbb{Q}_{N^{\alpha j}b}^{\theta\mathbf{1}} I_n^{\varphi}(Y_t) = \int_{\check{G}^n} d\mathbf{g}\, \varphi(\mathbf{g})\, \mathbf{Q}_{N^{\alpha j}b}^{\mathbf{g}} \theta^{|S_t|} \tag{57}$$

[recall that $S_t = \{i : \eta_i(t) \neq \dagger\}$].

If we take the limit as $j \to \infty$ in (57), then, by the construction of Section 3.1 and Corollary 14, we get

$$\theta \int_{\check{G}^n} d\mathbf{g}\, \varphi(\mathbf{g}) = \theta I_n^{\varphi}(\mathbf{1}).$$

On the other hand, if we take the limit as $j \to -\infty$ in (57), then we obtain

$$\theta^n \int_{\check{G}^n} d\mathbf{g}\, \varphi(\mathbf{g}) = I_n^{\varphi}(\theta\mathbf{1}).$$

Both claims then follow by Lemma 22.

(iv) From Lemma 8, we know that $\mathbb{Q}_b^{\theta\mathbf{1}}$-a.s. for l-a.e. $g \in G$ we have

$$Y_t(g) = \lim_{k\to\infty} N^k \int_{G_k} dh\, Y_t(g+h).$$

As Y_t is (spatially) stationary under $\mathbb{Q}_b^{\theta\mathbf{1}}$, the kth term in the sequence on the right-hand side has the same distribution as $I_1^{\varphi}\big((\Theta_{-k,0}Y)_t\big)$, where $\varphi = \mathbf{1}_{G_0}$. By part (iii),

$$\mathbb{Q}_b^{\theta\mathbf{1}} I_1^{\varphi}\big((\Theta_{-k,0}Y)_t\big)\big(1 - I_1^{\varphi}\big((\Theta_{-k,0}Y)_t\big)\big) \xrightarrow[k\to\infty]{} 0,$$

and so $\mathbb{Q}_b^{\theta\mathbf{1}}$-a.s. for l-a.e. $g \in G$ we have $Y_t(g) \in \{0,1\}$. $\square$

REFERENCES

[1] ARRATIA, R. (1982). Coalescing Brownian motions and the voter model on Z. Unpublished manuscript.

[2] BLUMENTHAL, R. M. and GETOOR, R. K. (1968). *Markov Processes and Potential Theory*. Academic Press, New York.

[3] BRAMSON, M., COX, J. T. and GRIFFEATH, D. (1986). Consolidation rates for two interacting systems in the plane. *Probab. Theory Related Fields* **73** 613–625.

[4] COX, J. T., FLEISCHMANN, K. and GREVEN, A. (1995). Comparison of interacting diffusions and application to their ergodic theory. Unpublished manuscript.

[5] COX, J. T. and GRIFFEATH, D. (1986). Diffusive clustering in the two dimensional voter model. *Ann. Probab.* **14** 347–370.

[6] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators, Part I: General Theory*. Interscience, New York.

[7] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.

[8] EVANS, S. N. (1988). Continuity properties of Gaussian stochastic processes indexed by a local field. *Proc. London Math. Soc.* **56** 380–416.
[9] EVANS, S. N. (1989). Local properties of Lévy processes on a totally disconnected group. *J. Theoret. Probab.* **2** 209–259.
[10] FLEISCHMANN, K. and GREVEN, A. (1994). Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Related Fields* **98** 517–566.
[11] FLEISCHMANN, K. and GREVEN, A. (1996). Time–space analysis of the cluster-formation in interacting diffusions. *Electron. J. Probab.* **1**.
[12] KLENKE, A. (1995). Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab.* **24** 660–697.
[13] KRENGEL, U. (1985). *Ergodic Theorems*. Walter de Gruyter, Berlin.
[14] MARCUS, M. B. and ROSEN, J. (1992). Sample path properties of the local times of strongly symmetric Markov proceses via Gaussian processes. *Ann. Probab.* **20** 1603–1684.
[15] MUELLER, C. and TRIBE, R. (1995). Stochastic PDE's arising from the long range contact and long range voter processes. *Probab. Theory Related Fields* **102** 519–545.
[16] SAWYER, S. (1976). Results for the stepping stone model for migration in population genetics. *Ann. Probab.* **4** 699–728.
[17] SAWYER, S. and FELSENSTEIN, J. (1983). Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.* **20** 1–10.
[18] SHIGA, T. (1980). An interacting system in population genetics. *J. Math. Kyoto Univ.* **20** 213–242.
[19] SHIGA, T. (1988). Stepping stone models in population genetics and population dynamics. In *Stochastic Processes in Physics and Engineering* (S. Albeverio, P. Blanchard, M. Hazewinkel and L. Streit, eds.) 345–355. Reidel, Dordrecht.
[20] SIMMONS, G. F. (1963). *Introduction to Topology and Modern Analysis*. McGraw-Hill, New York.
[21] VILENKIN, N. J. (1963). On a class of complete othonormal systems. *Amer. Math. Soc. Transl. Ser. 2* **28** 1–35.

DEPARTMENT OF STATISTICS # 3860
367 EVANS HALL
UNIVERSITY OF CALIFORNIA AT BERKELEY
BERKELEY, CALIFORNIA 94720-3860
E-MAIL: evans@stat.berkeley.edu

WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS (WIAS)
MOHRENSTRASSE 39
D-10117 BERLIN
GERMANY
E-MAIL: fleischmann@wias-berlin.de